

RECEIVED
CENTRAL FAX CENTER**APR 12 2006****CERTIFICATE OF TRANSMISSION****Date of Transmission: 12 April 2006**

I hereby certify that the following correspondence is being facsimile transmitted to the attention of the Director of the US Patent and Trademark Office on the above date via the following facsimile number: 571-273-8300.

PTO/SB/17 Fee Transmittal Form (1 sheet)
PTO-2038 Credit Card Payment Form (1 sheet)
Appeal Brief Under 37 C.F.R. §41.37 (10 sheets)
Appendix A (4 sheets)

Application Number 09/867,803
Confirmation No.: 6606
Filing Date: 31 May 2001
Document Submission Date: 12 April 2006

Art Unit: 2166
Examiner: Pham, Khanh B.
Inventor: Choi, Lawrence

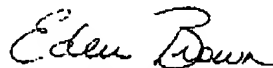
Docket: 1005-006**Pages: 17**

12 Apr 2006

Date

Eden Brown

Name of Certifier



Signature of Certifier

APR 12 2006

PTO/SB/17 (12-04)

Approved for use through 07/31/2006, OMB 0851-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

(Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.)

Effective on 12/08/2004. Fees pursuant to the Consolidated Appropriations Act, 2005 (H.R. 4818). FEE TRANSMITTAL For FY 2005		Complete if Known Application Number 09/867,803 Filing Date 31 May 2001 First Named Inventor Choi, Lawrence Examiner Name Pham, Khanh B. Art Unit 2188 Attorney Docket No. 1005-006	
<input checked="" type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27			
TOTAL AMOUNT OF PAYMENT	(\$)	250.00	

METHOD OF PAYMENT (check all that apply)

☐ Check ☒ Credit Card ☐ Money Order ☐ None ☐ Other (please identify): _____
☒ Deposit Account Deposit Account Number: 50-2504 Deposit Account Name: Michael N. Haynes
 For the above-identified deposit account, the Director is hereby authorized to: (check all that apply)
☐ Charge fee(s) indicated below ☐ Charge fee(s) indicated below, except for the filing fee
☒ Charge any additional fee(s) or underpayments of fee(s) under 37 CFR 1.16 and 1.17 ☒ Credit any overpayments

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

FEE CALCULATION**1. BASIC FILING, SEARCH, AND EXAMINATION FEES**

Application Type	FILING FEES		SEARCH FEES		EXAMINATION FEES		Fees Paid (\$)
	Fee (\$)	Small Entity Fee (\$)	Fee (\$)	Small Entity Fee (\$)	Fee (\$)	Small Entity Fee (\$)	
Utility	300	150	500	250	200	100	0
Design	200	100	100	50	130	65	
Plant	200	100	300	150	160	80	
Reissue	300	150	500	250	600	300	
Provisional	200	100	0	0	0	0	

2. EXCESS CLAIM FEES

Fee Description	Fee (\$)	Small Entity Fee (\$)
Each claim over 20 or, for Reissues, each claim over 20 and more than in the original patent	50	25
Each independent claim over 3 or, for Reissues, each independent claim more than in the original patent	200	100
Multiple dependent claims	360	180

Total Claims: _____ Extra Claims: 0 Fee (\$): 25 Fee Paid (\$): 0
 HP = highest number of total claims paid for, if greater than 20
 Indep. Claims: _____ Extra Claims: 0 Fee (\$): 100 Fee Paid (\$): 0
 HP = highest number of independent claims paid for, if greater than 3

3. APPLICATION SIZE FEE

If the specification and drawings exceed 100 sheets of paper, the application size fee due is \$250 (\$125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).

Total Sheets: _____ Extra Sheets: _____ Number of each additional 50 or fraction thereof: 0 Fee (\$): 125 Fee Paid (\$): 0
 - 100 = _____ / 50 = _____ (round up to a whole number) x _____

4. OTHER FEE(S)

	Fees Paid (\$)
Non-English Specification: \$130 fee (no small entity discount)	0
Other: Filing a Brief in Support of an Appeal	250

SUBMITTED BY		
Signature	<i>Michael N. Haynes</i>	Registration No. 40,014
Name (Print/Type)	Michael N. Haynes	Telephone 434-972-8988
		Date 12 Apr 2006

This collection of information is required by 37 CFR 1.136. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 30 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1460, Alexandria, VA 22312-1460. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1460, Alexandria, VA 22312-1460.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

APR 12 2006

PATENT

Application # 09/867,803

Attorney Docket # 1005-006

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s) : Lawrence J. Choi et al.
Application # : 09/867,803
Confirmation # : 6606
Filed : 31 May 2001
Application Title : METHOD AND SYSTEM FOR CLUSTERING
OPTIMIZATION AND APPLICATIONS
Art Unit # : 2167
Latest Examiner : Khanh B. Pham
Docket No. : 1005-006

Mail Stop Appeal Brief-Patents

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

APPEAL BRIEF UNDER 37 C.F.R. §41.37

Sir:

The Applicant respectfully submits this Appeal Brief in response to the Office Action of 5 October 2005 finally rejecting each of the pending claims 1, 5, and 6, and in furtherance of the Notice of Appeal filed 16 February 2006.

04/13/2006 MBINAS 00000037 09867803

01 FC:2402

250.00 0P

PATENT**Application # 09/867,803****Attorney Docket # 1005-006****I. REAL PARTY IN INTEREST**

The real party in interest is Rosetta Marketing Strategies Group, a corporation having a place of business at 103 Carnegie Center, Suite 202, Princeton, New Jersey 08540. *See* Assignment (recorded 31 May 2001 at reel 011881, frame 0782).

II. RELATED APPEALS AND INTERFERENCES

There are no related appeals or interferences.

III. STATUS OF CLAIMS

Claims 2, 7, and 8 were allowed. Claims 3 and 4 were objected to as being "dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims". Claims 1, 5, and 6 are pending in this application, have been finally rejected, and are the subject of this appeal. Each of claims 1, 5, and 6 are in independent form.

IV. STATUS OF AMENDMENTS

No amendments have been filed subsequent to the final rejections.

V. SUMMARY OF CLAIMED SUBJECT MATTER**Independent Claim 1**

Independent claim 1 recites a computer-assisted method for evaluating a cluster assignment for an observation (see at least pages 50-52; Fig. 4). The method comprises the activity of, for each of a plurality of observations, obtaining a data set containing no more than one proxy value for each of a plurality of variables, each variable having a plurality of possible values (see at least page 50; Fig. 4, element 4010), the data set also containing a cluster assignment for the observation, the cluster assignment identifying one cluster from a plurality of clusters (see at least page 50; Fig. 4, element 4010). The method further comprises the activity

PATENT**Application # 09/867,803****Attorney Docket # 1005-006**

of, for each observation from the plurality of observations, calculating a percent of proxy values for the plurality of variables that equals a mode of that observation's corresponding cluster's proxy values for the corresponding variables (see at least page 52; Fig. 4, element 4080). The method further comprises the activity of outputting the percent for each observation (see at least page 52; Fig. 4, element 4085).

Independent Claim 5

Independent claim 5 recites a computer-readable medium containing instructions for activities (see at least pages 50-52; Fig. 9, element 9300). The activities comprise, for each of a plurality of observations, obtaining a data set containing no more than one proxy value for each of a plurality of variables, each variable having a plurality of possible values (see at least page 50; Fig. 4, element 4010), the data set also containing a cluster assignment for the observation, the cluster assignment identifying one cluster from a plurality of clusters (see at least page 50; Fig. 4, element 4010). The activities further comprise, for each observation from the plurality of observations, calculating a percent of proxy values for the plurality of variables that equals a mode of that observation's corresponding cluster's proxy values for the corresponding variables (see at least page 52; Fig. 4, element 4080). The activities further comprise outputting the percent for each observation (see at least page 52; Fig. 4, element 4085).

Independent Claim 6

Independent claim 6 recites an apparatus for evaluating a cluster assignment for an observation (see at least pages 50-52; Fig. 9, Information Device 9). The apparatus comprises, for each of a plurality of observations, obtaining a data set containing no more than one proxy value for each of a plurality of variables, each variable having a plurality of possible values (see at least page 50; Fig. 4, element 4010), the data set also containing a cluster assignment for the observation, the cluster assignment identifying one cluster from a plurality of clusters (see at least page 50; Fig. 4, element 4010). The apparatus further comprises, for each observation from the plurality of observations, a means for calculating a percent of proxy values for the plurality of variables that equals a mode of that observation's corresponding cluster's proxy values for the

PATENT**Application # 09/867,803****Attorney Docket # 1005-006**

corresponding variables (see at least page 52; Fig. 4, element 4080). The apparatus further comprises a means for outputting the percent for each observation (see at least page 52; Fig. 4, element 4085).

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

Claims 1, 5, and 6 were rejected under 35 U.S.C. § 102(b) as being unpatentable over McLennan ("Statistics – A Powerful Edge", 1998).

VII. Argument – Claims 1, 5, and 6**A. General Legal Standards for Anticipation**

To anticipate expressly, the "invention must have been known to the art in the detail of the claim; that is, all of the elements and limitations of the claim must be shown in a single prior art reference, arranged as in the claim". *Karsten Mfg. Corp. v. Cleveland Golf Co.*, 242 F.3d 1376, 1383, 58 USPQ2d 1286, 1291 (Fed. Cir. 2001). The single reference must describe the claimed subject matter "with sufficient clarity and detail to establish that the subject matter existed in the prior art and that its existence was recognized by persons of ordinary skill in the field of the invention". *Crown Operations Int'l, LTD v. Solutia Inc.*, 289 F.3d 1367, 1375, 62 USPQ2d 1917, 1921 (Fed. Cir. 2002). Moreover, the prior art reference must be sufficient to enable one with ordinary skill in the art to practice the claimed invention. *In re Borst*, 345 F.2d 851, 855, 145 USPQ 554, 557 (CCPA 1965), *cert. denied*, 382 U.S. 973 (1966); *Amgen, Inc. v. Hoechst Marion Roussel, Inc.*, 314 F.3d 1313, 1354, 65 USPQ2d 1385, 1416 (Fed. Cir. 2003) ("A claimed invention cannot be anticipated by a prior art reference if the allegedly anticipatory disclosures cited as prior art are not enabled.") The USPTO "has the initial duty of supplying the factual basis for its rejection." *In re Warner*, 379 F.2d 1011, 154 USPQ 173, 178 (CCPA 1967).

B. The claimed "plurality of variables"

Each of claims 1, 5, and 6 recites, yet McLennan fails to teach or suggest, expressly or inherently, "for each of a plurality of observations", "obtaining a data set containing no more

PATENT**Application # 09/867,803****Attorney Docket # 1005-006**

than one proxy value for each of a **plurality of variables**, each variable having a plurality of possible values”.

The Final Office Action alleges, “in addition to battery’s [sic] file variables [sic], McLennan [sic] also teaches ‘battery size’ variable [sic], with the value ‘AA’ for each battery”. Thus, the Final Office Action apparently argues, leaning on page 81 of McLennan for support, that “battery size” is a variable. Yet the Final Office Action admits that McLennan describes only a single value, “AA”, for that alleged “battery size” variable, thus suggesting that McLennan at most discloses that “battery size” is a constant and not a variable. Because claims 1, 5, and 6 require that “each variable hav[e] a plurality of possible values”, McLennan’s “battery size” can not be a “variable” as that term is used in claims 1, 5, and 6.

To the extent that the Final Office Action alleges that McLennan discloses that each battery has “file variables”, Applicant respectfully requests citation to the precise location in McLennan supporting that allegation.

The only remaining alleged “variable” mentioned in the Final Office Action is “battery life”. Yet even assuming *arguendo* that McLennan does disclose that “battery life” is a “variable” (an assumption with which Applicant disagrees), McLennan nevertheless fails to teach or suggest the claimed “plurality of variables”.

Thus, McLennan fails to establish a *prima facie* case of anticipation since “all of the elements and limitations” of the claims are not present in McLennan. Accordingly, Applicant respectfully requests a reversal of each rejection of claims 1, 5, and 6.

C. The claimed “no more than one proxy value”

Each of claims 1, 5, and 6 recites, yet McLennan fails to teach or suggest, expressly or inherently, “for each of a plurality of observations”, “obtaining a data set containing **no more than one proxy value** for each of a plurality of variables, each variable having a plurality of possible values”.

The Final Office Action asserts that “McLennan teaches that each battery’s life variable [sic]... has only one value” and that “the data set contains one proxy value for each variables [sic] (i.e., minutes)”. Yet in describing McLennan’s “results” and “data”, McLennan states that

PATENT**Application # 09/867,803****Attorney Docket # 1005-006**

“[t]he lowest value is 363 and the highest value is 431”. Therefore, McLennan plainly indicates that the “data set” contains more than one “value” for “battery life”.

Thus, with regard to the alleged “battery life” variable, McLennan does not describe the claimed “no more than one proxy value”. Instead, to the extent that each alleged “value” for McLennan’s “battery life” is interpreted by the Final Office Action as equivalent to a “proxy value”, McLennan allegedly discloses “thirty” such alleged “proxy values”, which is much more than “one proxy value”.

Thus, McLennan fails to establish a *prima facie* case of anticipation since “all of the elements and limitations” of the claims are not present in McLennan. Accordingly, Applicant respectfully requests a reversal of each rejection of claims 1, 5, and 6.

D. The claimed “percent of proxy values for the plurality of variables”

Each of claims 1, 5, and 6 recites, yet McLennan fails to teach or suggest, expressly or inherently, “calculating a percent of proxy values for the plurality of variables that equals a mode of that observation’s corresponding cluster’s proxy values for the corresponding variables”.

As described above, McLennan fails to teach or suggest a “plurality of variables”. Moreover, McLennan fails to teach or suggest, “calculating a percent of proxy values for the plurality of variables”. To the extent the Final Office Action alleges that McLennan discloses, “calculating a percent of proxy values” for “battery life”, McLennan nevertheless fails to teach or suggest, “calculating a percent of proxy values” for any other alleged “variable”. Thus, McLennan fails to teach or suggest, “calculating... for the plurality of variables”.

Thus, McLennan fails to establish a *prima facie* case of anticipation since “all of the elements and limitations” of the claims are not present in McLennan. Accordingly, Applicant respectfully requests a reversal of each rejection of claims 1, 5, and 6.

E. The claimed “plurality of clusters”

Each of claims 1, 5, and 6 recites, yet McLennan fails to teach or suggest, expressly or inherently, “for each of a plurality of observations, obtaining a data set ... containing a cluster

PATENT**Application # 09/867,803****Attorney Docket # 1005-006**

assignment for the observation, the cluster assignment identifying one cluster from a **plurality of clusters**"; and "for each observation from the plurality of observations", "calculating a percent of proxy values for the plurality of variables that equals a mode of that observation's corresponding cluster's proxy values for the corresponding variables".

The Final Office Action asserts that "McLennan clearly teaches at page 83, 'a mode of [sic] corresponding cluster's proxy values', wherein the cluster (i.e., 'class interval') is in the range 360-439, and that the "table at page 83 shows the mode of the observations [sic] corresponding cluster is 7 (i.e., 'the most frequently observed value')".

To the extent Applicant can interpret them, these assertions at least ignore and mischaracterize the claim language. Moreover, it appears that these assertions prove that McLennan does not anticipate claims 1, 5 and 6.

McLennan shows on page 81 that in the "class interval" of "390-399", the tally was seven tick marks out of thirty total, and thus the "frequency" was "7" out of a "total" of "30". If, as asserted by the Final Office Action, McLennan's disclosed "7" is the "the mode of the observations [sic] corresponding cluster", ostensibly, as asserted by McLennan on page 144, "because this value occurred the most often", then the Final Office Action is apparently arguing that McLennan discloses only a single "cluster" which contains all 30 "battery life" measurements, for which "7" is the alleged "mode".

Yet the claims require that "the cluster assignment identify[] one cluster from a **plurality of clusters**". So where does McLennan teach or suggest a "plurality of clusters"? Applicant respectfully submits that the applied portions of McLennan do not teach or suggest a "plurality of clusters".

Thus, McLennan fails to establish a *prima facie* case of anticipation since "all of the elements and limitations" of the claims are not present in McLennan. Accordingly, Applicant respectfully requests a reversal of each rejection of claims 1, 5, and 6.

F. The claimed "mode of" an "observation's" "cluster's proxy values"

Each of claims 1, 5, and 6 recites, yet McLennan fails to teach or suggest, expressly or inherently, "calculating a percent of proxy values for the plurality of variables that equals a

PATENT**Application # 09/867,803****Attorney Docket # 1005-006**

mode of that observation's corresponding cluster's proxy values for the corresponding variables".

As discussed above, McLennan fails to teach or suggest more than one "variable", and thus does not teach or suggest the claimed "variables".

But assuming *arguendo* that McLennan does disclose even a single variable (an assumption with which Applicant disagrees), which the Final Office Action asserts is "battery life", the Final Office Action admits that McLennan does not teach or suggest more than one "proxy value" for even that single variable, let alone a "plurality of variables".

But assuming *arguendo* that McLennan does disclose such "proxy values" (an assumption with which Applicant disagrees), as explained above, McLennan fails to teach or suggest "a plurality of clusters", and instead at most discloses only a single "cluster".

Thus, to the extent that the Final Office Action interprets McLennan's "a most frequently observed value" as a "mode", that alleged "mode" is for all the "battery life" "values" rather than being "a mode of" a single "cluster's proxy values", that single cluster being one of "a plurality of clusters".

Applicant respectfully submits that McLennan fails to teach or suggest "a mode of that observation's corresponding cluster's proxy values" as recited in each of claims 1, 5, and 6. Thus, McLennan fails to establish a *prima facie* case of anticipation since "all of the elements and limitations" of the claims are not present in McLennan. Accordingly, Applicant respectfully requests a reversal of each rejection of claims 1, 5, and 6.

G. The claimed "obtaining a data set" for each observation

Each of claims 1, 5, and 6 recites, yet McLennan fails to teach or suggest, expressly or inherently, "**for each of a plurality of observations**", "**obtaining a data set** containing no more than one proxy value for each of a plurality of variables, each variable having a plurality of possible values, the data set also containing a cluster assignment for the observation, the cluster assignment identifying one cluster from a plurality of clusters"; and "for each observation from the plurality of observations, calculating a percent of proxy values for the plurality of variables that equals a mode of that observation's corresponding cluster's proxy values for the

PATENT**Application # 09/867,803****Attorney Docket # 1005-006**

corresponding variables”.

The Final Office Action asserts that “McLennan teaches the data set obtained from each of a plurality of observations of battery life for thirty AA size batteries, the data set contains one proxy value for each variables [sic] (i.e., the minutes), wherein the possible value for each variables [sic] ranges from 360 to 439”. Thus, the Final Office Action appears to assert that each entry in the list of numbers presented near the top of McLennan’s page 81, is both a “proxy value for each of a plurality of variables” and an “observation”.

Yet claims 1, 5, and 6 do not recite a “data set obtained from... a plurality of observations”, but instead recite, “obtaining a data set” “for each ... observation []”. Thus, once again, the Final Office Action has mischaracterized and misinterpreted the claimed subject matter. If McLennan’s first “battery life” “value” and “observation” is 423, where does McLennan disclose the “data set” “for” that alleged “observation”?

Thus, McLennan fails to establish a *prima facie* case of anticipation since “all of the elements and limitations” of the claims are not present in McLennan. Accordingly, Applicant respectfully requests a reversal of each rejection of claims 1, 5, and 6.

VIII. CLAIMS APPENDIX

Appendix A sets forth all pending claims in the state in which they were appealed.

IX. EVIDENCE APPENDICES

There are no evidence appendices.

X. RELATED PROCEEDINGS APPENDIX

There are no related proceedings.

PATENT

Application # 09/867,803

Attorney Docket # 1005-006

SUMMARY

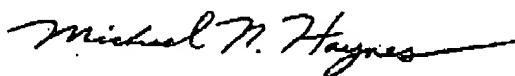
In view of the above, Applicant submits that all claims on appeal distinguish over the applied art and respectfully requests that the rejections of these claims should be reversed.

Applicant therefore respectfully requests that the Board of Patent Appeals and Interferences reverse the decision rejecting claims 92-110 and direct that the application be passed to issue.

The Office is hereby authorized to charge any additional fees or credit any overpayments under 37 C.F.R. 1.16 or 1.17 to Deposit Account No. 50-2504. The Examiner is invited to contact the undersigned at 434-972-9988 to discuss any matter regarding this application.

Respectfully submitted,

Michael Haynes PLC



Michael N. Haynes
Registration No. 40,014

Date: 11 April 2006

1341 Huntersfield Close
Keswick, VA 22947
Telephone: 434-972-9988
Facsimile: 815-550-8850

APR 12 2006

PATENT

Serial No. 09/867,803

Attorney Docket No. 1005-006

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s) : Lawrence J. Choi et al.
Serial No. : 09/867,803
Filed : 31 May 2001
For : METHOD AND SYSTEM FOR CLUSTERING
OPTIMIZATION AND APPLICATIONS
Art Unit : 2177
Examiner : Khanh B. Pham

Mail Stop Appeal Brief-Patents

Commissioner for Patents

P.O. Box 1450

Alexandria, VA 22313-1450

APPENDIX A

1. A computer-assisted method for evaluating a cluster assignment for an observation, comprising the activities of:

for each of a plurality of observations, obtaining a data set containing no more than one proxy value for each of a plurality of variables, each variable having a plurality of possible values, the data set also containing a cluster assignment for the observation, the cluster assignment identifying one cluster from a plurality of clusters;

for each observation from the plurality of observations, calculating a percent of proxy values for the plurality of variables that equals a mode of that observation's corresponding cluster's proxy values for the corresponding variables; and
outputting the percent for each observation.

PATENT**Serial No. 09/867,803****Attorney Docket No. 1005-006**

2. A computer-assisted method for evaluating a cluster assignment for an observation, comprising the activities of:
 - for each of a plurality of observations, obtaining a data set containing no more than one proxy value for each of a plurality of variables, each variable having a plurality of possible values, the data set also containing a cluster assignment for the observation;
 - for each observation from the plurality of observations, estimating a purposeful probability that a particular possible value from the plurality of possible values for a particular variable will be purposefully provided by observations assigned to a particular cluster from a plurality of clusters; and
 - outputting each purposeful probability.
3. The method of claim 1, further comprising the activities of:
 - for each observation from the plurality of observations in each cluster from the plurality of clusters, calculating a serendipity probability for each possible value, the serendipity probability is a measure of a probability that an observation in a particular cluster will be randomly associated with any one of the plurality of possible values for a particular variable;
 - for each observation from the plurality of observations, calculating a ratio of the purposeful probability to the serendipity probability;
 - for each observation from the plurality of observations, calculating a logarithm of the ratio to obtain composition analysis score; and
 - outputting the composition analysis scores for each observation in each cluster.
4. The method of claim 1, further comprising the activities of:
 - for each observation from the plurality of observations, assuming that before the observation can be made, the observation has an equal probability of being in any identified cluster from the plurality of clusters;
 - for each observation from the plurality of observations, assuming that the purposeful probabilities are true;

PATENT**Serial No. 09/867,803****Attorney Docket No. 1005-006**

for each observation from the plurality of observations, using Bayes' Theorem to calculate a Bayes probability that a particular observation can be in each cluster conditional upon the observation's proxy value to each variable;

outputting the Bayes probability that each observation can be in each cluster.

5. A computer-readable medium containing instructions for activities comprising:

for each of a plurality of observations, obtaining a data set containing no more than one proxy value for each of a plurality of variables, each variable having a plurality of possible values, the data set also containing a cluster assignment for the observation, the cluster assignment identifying one cluster from a plurality of clusters;

for each observation from the plurality of observations, calculating a percent of proxy values for the plurality of variables that equals a mode of that observation's corresponding cluster's proxy values for the corresponding variables; and

outputting the percent for each observation.

6. An apparatus for evaluating a cluster assignment for an observation, comprising:

for each of a plurality of observations, means for obtaining a data set containing no more than one proxy value for each of a plurality of variables, each variable having a plurality of possible values, the data set also containing a cluster assignment for the observation, the cluster assignment identifying one cluster from a plurality of clusters;

for each observation from the plurality of observations, means for calculating a percent of proxy values for the plurality of variables that equals a mode of that observation's corresponding cluster's proxy values for the corresponding variables; and

means for outputting the percent for each observation.

7. A computer-readable medium containing instructions for activities comprising:

for each of a plurality of observations, obtaining a data set containing no more than one proxy value for each of a plurality of variables, each variable having a plurality of possible values, the data set also containing a cluster assignment for the observation;

PATENT**Serial No. 09/867,803****Attorney Docket No. 1005-006**

for each observation from the plurality of observations, estimating a purposeful probability that a particular possible value from the plurality of possible values for a particular variable will be purposefully provided by observations assigned to a particular cluster from a plurality of clusters; and

outputting each purposeful probability.

8. An apparatus for evaluating a cluster assignment for an observation, comprising:

for each of a plurality of observations, means for obtaining a data set containing no more than one proxy value for each of a plurality of variables, each variable having a plurality of possible values, the data set also containing a cluster assignment for the observation;

for each observation from the plurality of observations, means for estimating a purposeful probability that a particular possible value from the plurality of possible values for a particular variable will be purposefully provided by observations assigned to a particular cluster from a plurality of clusters; and

means for outputting each purposeful probability.